Paths to accuracy for radiation parameterizations in atmospheric models

Robert Pincus[1] and Bjorn Stevens[2]

————

Robert Pincus, 325 Broadway R/PSD1, Boulder, CO, 80305, USA. (Robert.Pincus@colorado.edu)

Bjorn Stevens, Max Planck Institute for Meteorology, Bundesstraße 53, D-20146 Hamburg, Germany. (Bjorn.Stevens@mpimet.mpg.de)

[1]University of Colorado and NOAA Earth System Research Lab, Physical Sciences Division, Boulder, Colorado, USA.

[2]Max Planck Institute for Meteorology, Hamburg, Germany

Abstract.   Radiative transfer is sufficiently well-understood that its parameterization in atmospheric models is primarily an effort to balance computational cost and accuracy. The most common approach is to compute radiative transfer with the highest practical spectral accuracy but infrequently in time and/or space, though errors introduced by this approximation are difficult to quantify. An alternative is to perform spectrally sparse calculations frequently in time using randomly-chosen spectral quadrature points. Here we show that purely random quadrature points, though effective in some large-eddy simulations, are not a good choice for models in which the land surface responds to radiative fluxes because surface temperature perturbations can be large enough, and persistent long enough, to affect model evolution. These errors may be mitigated by choosing teams of spectral points designed to limit the maximum surface flux error; teams, rather than individual quadrature points, are then chosen randomly. The approach is implemented in the ECHAM6 global model and the results are examined using "perfect-model" experiments on timescales ranging from a day to a month. In this application the approach introduces errors commensurate with the infrequent calculation of broadband calculations for the same computational cost. But because teams need not increase with size, and indeed may become better and more balanced with increased spectral density, improvements in radiative transfer may not need to be traded off against spatio-temporal sampling.

1. What does it mean for a parameterization to be accurate?

Unlike many closure problems faced in models of the atmosphere, the environmental factors that control the distribution of radiation in the atmosphere are very well understood, so the solution to fully specified problems is known to great accuracy. Radiation parameterizations therefore seek primarily to find an acceptable compromise between accuracy and computational cost. The accuracy of radiative transfer calculations may be measured via comparison to benchmark models [*Oreopoulos et al.*, 2012] which are themselves known to be in excellent agreement with observations [*Mlawer et al.*, 2000; *Turner et al.*, 2004]. Comparisons are normally made for clear-sky conditions, consistent with the way the parameterizations of absorption by gases are developed.

State-of-the-art radiation parameterizations can reproduce benchmark calculations to within 1% for shortwave fluxes and fractions of a percent for longwave fluxes [*Oreopoulos et al.*, 2012] but this accuracy is so computationally expensive that radiation parameterizations cannot be applied at every timestep of the model. Instead, radiative heating and cooling rates are normally updated less frequently than are model dynamics and, in most cases, other physical parameterizations. The choice to update radiative heating rates less frequently than other fields is an approximation made, not in the radiation parameterization, but in the coupling to the rest of the model. The simulation errors caused by this approximation may range from modest changes in temperature fields [*Xu and Randall*, 1995; *Morcrette*, 2000] to the introduction of more dramatic instabilities [*Pauluis and Emanuel*, 2004] but are generally difficult to quantify. To minimize simulation errors

45 prudence dictates that the radiation time-step be as close to the dynamical timestep as

46 can be afforded, although precisely how close is a subjective choice.

47   Several approaches have been developed to accelerate the calculation of radiative fluxes

48 to allow for more frequent calculation. One is to use physically-based radiative transfer

49 models to train fast statistical models (normally artificial neural networks) to emulate

50 fluxes based on the state of the atmosphere [e.g. *Chevallier et al.*, 1998; *Krasnopolsky et al.*,

51 2008]. An intermediate tactic is to apply physical models sparsely in space and/or time,

52 use simple statistical models (e.g. regression) to predict changes since the last radiation

53 time step, and selectively update calculations based on the expected error [*Venema et al.*,

54 2007]. A third alternative exploits two facts – that cloud properties vary much more

55 quickly in the atmosphere than does the concentration of gases, and that variations in

56 clouds and gases affect fluxes in different, roughly disjoint spectral regions – to motivate

57 updating only the cloud-affected portions of the spectrum at high frequency [*Manners*

58 *et al.*, 2009]. (*Räisänen and Barker* [2004] and *Hill et al.* [2011] take a similar approach

59 to the related problem of representing cloud variability.) The calculation of cloud-affected

60 fluxes can be further accelerated by reducing the spectral detail used to treat absorption

61 by gases [*Manners et al.*, 2009].

62   Each of these methods, including infrequent radiation calculations, represent approxi-

63 mations which introduce errors in radiative heating rates. These errors depend on many

64 factors including how quickly the optical properties of the atmosphere are changing. But

65 the error characteristics of an approximation can be crucially important in determining

66 whether the approximation affects model evolution. Since radiative fluxes at the top of

67 the atmosphere are essentially in balance (after accounting for ocean heat storage), for

68 example, even small (< 1 W/m^2) biases in radiative fluxes affect multi-decadal simula-

69 tions and must be "tuned" away [*Mauritsen et al.*, 2012] and/or balanced by compensating

70 errors. Random, uncorrelated noise, on the other hand, does not affect the statistical evo-

71 lution of most models, whether that noise comes from parameterizations of gravity wave

72 drag [*Eckermann*, 2012; *Lott et al.*, 2012] or radiation [*Pincus et al.*, 2003] or is externally

73 applied in an effort to diversify ensembles of medium-range forecasts [*Buizza et al.*, 1999].

74 For the purposes of parameterization development this implies that unbiased algorithms,

75 even if they introduce even quite substantial noise in heating rates, can be more accurate,

76 in the sense of introducing smaller changes in model evolution, than other approximations

77 including detailed algorithms used infrequently.

78   Here we describe an approach to radiative transfer parameterization that emphasizes

79 the accuracy relevant for hydrodynamic models, including both the radiation calculations

80 and the ways those calculations are coupled to the rest of the model. The approach takes

81 advantage of the local homogenization of heating rates arising from small scale fluid dy-

82 namical processes. We have implemented these ideas in a new radiation package, PSrad

83 (named because it is a postscript to the RRTMG package from which it descends), and

84 initially implemented in ECHAM climate model. PSrad is unique in that only a small sam-

85 ple of the full broadband spectral integration is performed but these calculations are per-

86 formed at each time step. This spectral sampling introduces grid-scale noise in radiative

87 fluxes, as does the more common use of stochastic samples to represent the subgrid-scale

88 distribution of cloud properties [*Pincus et al.*, 2003]. Experiments show that ECHAM

89 is insensitive to even large grid-scale perturbations to radiative heating rates within the

90 atmosphere, but that significant perturbations in surface fluxes can introduce systematic

91 biases in the model trajectory. Simulation bias can be limited by bounding errors in sur-

92 face fluxes using carefully selected subsets of the broadband calculation. The approach

93 is applicable to dynamical models at all scales even as significant noise is introduced into

94 individual calculations.

95   The next section details several strategies for coupling radiation calculations to model

96 integrations, including infrequent radiation calculations and spectral subsets computed

97 at higher frequency. Section 3 describes the code we have developed to implement these

98 strategies and section 4 the impact of two classes of approximations on forecasting and

99 quasi-climatological time scales in climate model integrations. Section 5 discusses impli-

100 cations for weather forecasting and climate projection applications.

2. Strategies for spectral integration

Models of the atmosphere require broadband radiation calculations, i.e. those that

account for all wavelengths of radiation emitted by the sun or the earth and its atmosphere.

In parameterizations of radiative transfer this spectral integration is accomplished using

weighted sums

$$F(x,y,z,t) = \sum_{g}^{G} w_g F_g(x,y,z,t) \tag{1}$$

101 where the individual fluxes F_g are computed using optical properties and boundary condi-

102 tions appropriate to each pseudo-spectral interval (quadrature point). These quadrature

103 points are frequently determined using k-distributions [*Fu and Liou*, 1992; *Lacis and*

104 *Oinas*, 1991]; following this nomenclature we refer to these intervals as "g-points". In the

105 shortwave the weights w_g account for the distribution of incoming solar energy among

106 the spectral intervals while in the longwave the spectral distribution of flux, and so the

107 weights, depends on temperature.

In most models of the atmosphere radiative fields are updated less frequently then other

variables, i.e.

$$F(x, y, z, t) \approx c(x, y, t) \sum_{g}^{G} w_g F_g(x, y, z, t_{\text{rad}}(t) \mid t_{\text{rad}}(t) - t \leq \Delta t_{\text{rad}}) \qquad (2)$$

108 where c represents correction factors that may be applied to account for time-varying

109 solar zenith angles, surface temperatures, etc. In some implementations [e.g. *Morcrette*,

110 2000] spatial resolution may also be reduced.

2.1. Monte Carlo Spectral Integration

Infrequent broadband calculations (e.g. Equation 2) can be described as a "spectrally

dense, temporally sparse" approach to computing radiation transfer. Monte Carlo Spectral

Integration [MCSI; see *Pincus and Stevens*, 2009] reverses these densities:

$$F(x, y, z, t) \approx G/G' \sum_{g \in [1..G]}^{G'} w_g F_g(x, y, z, t) \qquad (3)$$

111 where each of the G' samples is chosen randomly with replacement at each location and

112 time step.

113 MCSI was initially introduced for large eddy simulation where it has the advantages of

114 being consistent (i.e. converging with increasing temporal and/or spatial resolution, as

115 do the other approximations used in LES) and in explicitly sampling temporal variability,

116 especially in cloud optical properties. It takes advantage of the fact that local fluid

117 instabilities homogenize sampling noise on small scales while, on the larger spatial and

118 temporal scales where heating rates can effect the overall evolution of the flow, sampling

119 noise is small. In several applications, including purely radiatively-driven flow, LES using

120 MCSI is statistically indistinguishable from LES using benchmark radiation calculations

121 (e.g. Eq. 1) [*Pincus and Stevens*, 2009].

122   MCSI implemented similarly in a global model introduces much larger and more sys-

123 tematic errors. The green line in Fig. 1 shows the global root-mean-square difference in

124 2-m air temperature as a function of forecast lead time between a reference calculation

125 that computes broadband radiation at every time step and every grid-point and one using

126 the same radiation code (described in Sec. 3) but applying Eq. 3 with $G' = 1$. The

127 figure shows the average over 29 independent forecasts (see Sec. 4). RMS differences with

128 respect to the reference forecast exceed 1.5 K after the first day and grow over time. A

129 ~ 0.5 K diurnal cycle tracks the diurnal variation in global mean 2-m temperature and

130 occurs because the land surface is not homogeneously distributed over Earth's surface.

131   Why is the MCSI approximation accurate (in the sense of not disturbing the flow) in

132 large eddy simulations but not in a global model? There are at least two significant dis-

133 tinctions. First, the parameterizations used in global models, especially those for deep

134 and shallow convection, depend more non-linearly on the atmospheric state than do the

135 simple sub-grid scale models for turbulence and microphysics used in LES, so even unbi-

136 ased random sampling noise can bias the flow through non-linearities. More importantly,

137 surface properties are frequently fixed in large-eddy simulations, while global models al-

138 most always include land surfaces whose temperature changes in response to surface fluxes.

139 Perturbations to the surface temperature caused by sampling errors are not homogeneized

140 by mixing with neighboring columns, while errors within the atmosphere are mixed by the

141 fluid flow. The impact of surface temperature perturbations dominates: on an aquaplanet

142 with globally specified sea surface temperatures (Fig. 1, purple line), differences between

143 simulations using Eq. 3 and simulations using Eq. 1 are small.

2.2. Bounding errors in surface fluxes using teams of spectral points

144 Figure 1 implies that the magnitude of instantaneous surface flux errors must be

145 bounded if a radiation parameterization is to be useful in models including land sur-

146 faces. One approach would be to simply increase the numbers of samples, G', chosen at

147 each time step but this is painfully slow: like all Monte Carlo estimates the root-mean-

148 square error, for example, decreases as $1/\sqrt{G'}$ [see, e.g., Sec. 4.2.5 of *Evans and Marshak*,

149 2005]. An more efficient strategy is to generate sets of g-point "teams" constructed to

150 minimize some measure of sampling error, and to sample these teams randomly.

151 Assume a set of A representative atmospheres and define a cost function C as some

152 measure of the error accumulated over L possible estimates F_l^a of the true flux F^a in each

153 atmosphere. The number of teams, M, may be chosen to be small divisor of the total

154 number of g-points G so that each team has the same number $m = G/M$ of quadrature

155 points. The members of these teams may then be chosen to minimize C.

156 We have computed teams for several values of M using the radiative transfer code de-

157 scribed in Sec. 3. Our set of representative atmospheres is obtained from four snapshots

158 taken over the course of a single day from a free run of ECHAM6 ($A \approx 73000$). We opti-

159 mize over $L = MA$ clear-sky fluxes consistent with the way k-distributions are normally

160 constructed. Since our goal is to minimize the possibility of very large errors in surface

161 fluxes we use the 95% error in surface fluxes E_{srf}^{95} as our cost function C. Our minimiza-

162 tion of C is informal: we compute fluxes for each g-point individually and choose the

163 M g-points which make the worst proxies as that first member of each teams. For each

164 remaining team member we process teams randomly and choose the remaining g-point

165 that minimizes the cost function for the provisional team. It is likely that the balance of

166 teams could be modestly improved though further optimization (by simulated annealing,

167 for example).

Fluxes can then be computed by choosing one of the teams at random:

$$F(x, y, z, t) \approx M \sum_{g(i \in [1..M])}^{m} w_g F_g(x, y, z, t) \tag{4}$$

168 Because each g-point is included in exactly one team Eq. 4, like Eq. 3, is an unbiased

169 estimate of the true flux given by Eq. 1.

170 Using teams of g-points is effective in limiting the error for a given computational cost

171 (Figure 2). Though the teams are chosen to minimize errors in clear-sky surface fluxes

172 the analogous errors for all skies are commensurate, slightly lower in the shortwave where

173 the presence of clouds simply reduces downwelling flux and slightly larger in the longwave

174 where clouds may change the spectral distribution of flux.

175 Teams constructed in this way are more efficient in reducing error as team size increases.

176 Figure 3 shows the ratio of errors obtained using teams of a given size (Eq. 4) to those

177 using Monte Carlo samples (Eq. 3). The ratio is small in both the shortwave and longwave

178 for $m = 2$ but increases at four or more samples, such that random sampling using Eq.

179 3 can achieve commensurate accuracy only by increasing sample sizes by a factor of ten

180 or greater - in other words, by increasing computational cost to nearly that of broadband

181 integration using Eq. 1.

3. PSrad/RRTMG: A new radiation code for climate models

182 We have developed a new radiation package, PSrad/RRTMG (so named because it is a

183 postscript to the RRTMG package), designed for use in models of the atmosphere. The

184 longwave and shortwave components are organized along parallel lines: driver modules

185 call routines to compute the optical properties of gases, aerosols, and clouds, then combine

186 these to compute fluxes at the boundaries and the interfaces between model layers. The

187 codes are modeled after the RRTMG package [*Mlawer et al.*, 1997; *Iacono et al.*, 2008]

188 and use the k-distribution from this package to determine gas optical thickness from

189 concentrations, temperature, and humidity. (These k-distributions are well-validated and

190 among the most accurate available; see *Oreopoulos et al.* [2012].) We follow the original

191 RRTMG codes in using the two-stream approximation [after *Meador and Weaver*, 1980]

192 to compute layer reflectance and transmittance and adding [after *Oreopoulos and Barker*,

193 1999] to compute flux profiles in the shortwave; we use the linear-in-tau approximation

194 for the thermal source function [*Mlawer et al.*, 1997] and consider only emission and

195 absorption in the longwave. Cloud and aerosol optical properties are determined from

196 custom-built lookup tables [*Kinne et al.*, 2012].

197 Sub-grid scale variability is treated using "sub-columns" [see, e.g. *Räisänen et al.*, 2004;

198 *Pincus et al.*, 2006]: discrete random samples, each treated as internally homogeneous,

199 that are consistent with the distributions of possible cloud states within each column, in-

200 cluding fractional cloudiness in each layer and assumptions about the vertical correlations

201 between layers ("cloud overlap"). This treatment is a generalization of the Monte Carlo

202 Independent Pixel Approximation [*Pincus et al.*, 2003] and may be further generalized to

203 include other kinds of variability, including the distribution of cloud liquid or ice water

204 content as implied by, for example, the *Tompkins* [2002] cloud scheme.

205 PSrad supports a range of choices for spectral sampling, including broadband integration

206 (all G quadrature points in order), arbitrary numbers of randomly chosen quadrature

207 points for application of Eq. 3, and a finite number of pre-determined "leagues" of g-point

208 teams as described in Sec. 2.

209 Though PSrad is currently intended as a drop-in replacement for RRTMG it was im-

210 plemented almost entirely from scratch (of the original code, only the subroutine that

211 computes longwave gas optical properties remains). The most important technical dif-

212 ference lies in the organization: each subroutine is vectorized over model columns, which

213 increases computational efficiency on a wide range of platforms. Operational centers

214 such as the European Centre for Medium-Range Weather Forecasts have often modified

215 RRTMG in this way [*Morcrette et al.*, 2008].

4. Assessing approximation impacts in a global model

216 We have implemented PSrad in ECHAM6 [*Stevens et al.*, 2012], a state-of-the-art at-

217 mospheric model used for climate simulations. We perform simulations with a version of

218 the model differing modestly from the version used to produce data for the Fifth Coupled

219 Model Intercomparison Project [*Taylor et al.*, 2012]. The model is run at a horizontal

220 resolution of T63 using 47 levels that extend to 1 hPa. These experiments use a 7.5 min

221 time step.

222 We consider an ensemble of 29 month-long integrations starting from initial conditions

223 valid at 0Z on 1 Apr of the years 1976-2004 as simulated by the model in a long integration

224 using specified, time-varying sea-surface temperatures. The benchmark is an integration

225 in which radiation fields are calculated just as frequently as the tendencies from other

226 physical parameterizations (i.e the radiation and physics time steps are the same). We

227 use this reference forecast to assess the error introduced by increasing the interval between

228 broadband radiation calculations, on the one hand, and by limiting the number of spectral

229 quadrature points used at each time, on the other.

230 Root-mean-square differences with the reference forecast grow with time (Figure 4),

231 but can be divided into roughly three regimes: slow but accelerating error growth in the

232 first ten days, rapid error growth over the next ten days, and roughly saturated errors

233 in the last ten days. This may be very loosely described as the transition from weather

234 (where individual trajectories are followed) to climate (where the statistics of trajectories

235 are of interest), which leads us to examine errors in the first ten days as one might

236 evaluate forecasts, but to assess errors in the last ten days statistically, as one might

237 assess climatologies.

238 The Monte Carlo sampling of fractional cloudiness and overlap introduces noise into

239 the fluxes and causes even a second independent reference forecast to diverge from the

240 benchmark over time: RMS differences between two sets of reference forecasts rise from

241 about 0.05 K during the first day to almost 1.5 K after ten days (Figure 5, top panel,

242 purple line), which we take as the rough limit of deterministic forecasts. Increasing the

243 sparsity of radiation calculations in either time (purple lines, $\Delta t_{\mathrm{rad}} \in [15\mathrm{min}, 1\mathrm{h}, 2\mathrm{h}, 3\mathrm{h}]$)

244 or spectral quadrature points (green lines, $m_{\mathrm{LW}} \in [4, 10]$ and $m_{\mathrm{SW}} \in [7, 16]$) increases the

245 error by modest amounts in an absolute sense. No evidence has been found that such ap-

246 proximation trigger dramatic changes in the simulation [c.f. *Pauluis and Emanuel*, 2004],

247 though even hourly radiation calls increases the RMS difference by 50% of the difference

248 introduced by sampling noise (Figure 5, bottom panel). The mean RMS difference over

249 the first ten days (Figure 6) is quite tightly related to computational cost (here expressed

250 as the number of shortwave radiation computations per day).

251   On timescales where the climatology of the model dominates the character of the solu-

252 tion it is more informative to assess the degree to which each approximation produces a

253 statistical distribution of temperatures consistent with the reference forecast. We apply

254 the two-sided Student's t-test, for each approximation at every grid point and every time

255 step during days 21-31, to compute the likelihood (p-value) that the distribution of 2-m

256 temperatures across the 29 ensemble members is statistically indistinguishable at the 95%

257 level between the reference forecast and the forecast using the approximation. Because we

258 perform so many t-tests ($\sim$ 18,500 per time step) roughly 5% have p-values corresponding

259 to "significant" differences (at the 95% level) even if the underlying distribution of 2-m

260 temperatures is the same in both experiments. This may be controlled for using false

261 discovery rate estimation [e.g. *Wilks*, 2006] which exploits the known distribution of p-

262 values expected under the null hypothesis to estimate η_0, the proportion of uninteresting

263 (or truly insignificant) p-values [*Strimmer*, 2008a], at every time step.

264   Independent realizations of the reference run (i.e. two runs making broadband radiation

265 calculations at every time step, but using different random number sequences to sample

266 cloud states with McICA) are statistically indistinguishable: the time-mean value of η_0

267 computed from this pair of experiments is 1. This is almost but not quite true when

268 comparing the reference run to any of the possible approximations (see Table 1). One

269 interpretation of $\overline{\eta_0}$ is as the average fraction of the planet over which a given approxi-

270 mation does not change the simulation significantly. This fraction is greater than 94%

271 for all approximations except MCSI, indicating that changes are detectable but mod-

272 est. The test statistic is not entirely robust: according to this measure radiation time

273 steps of $T = 1\mathrm{h} = 8\Delta t$ are slightly more consistent with reference calculations than are

274 $T = 15\mathrm{min} = 2\Delta t$, which is physically implausible. Given this caveat we note that, in this

275 relatively coarse-resolution model, infrequent broadband radiation calculations introduce

276 slightly smaller changes compared to frequent calls using spectral teams with the same

277 computational cost (c.f. lines 2 and 6 of Table 1).

5. Conclusions: Parameterization error, simulation error, and the coupling of radiative transfer to atmospheric models

278   Radiative fluxes respond nearly instantaneously to changes in the optical properties of

279 the atmosphere, so the parameterization of radiation is normally considered a "one-way"

280 problem in which the model provides the state of the atmosphere and the parameterization

281 computes the heating rates and boundary fluxes. In the absence of coupling between

282 radiation and model dynamics one naturally seeks instantaneous radiation calculations

283 that are as accurate as is computationally feasible. Focusing on the accuracy of the overall

284 simulation, including the way radiation calculations are coupled to the model, may allow

285 for other kinds of optimization. As one example, the k-distribution developed for RRTMG

286 was designed, as are most parameterizations, to minimize broadband flux errors. RRTMG

287 aims to balance accuracy with computational cost primarily by minimizing the number

288 of g-points. It would be possible, however, to construct k-distributions using different

289 cost functions, and distributions constructed to minimize errors both across and among

290 teams might be able to achieve greater aggregate accuracy by using more g-points while

291 still limiting noise in surface fluxes. This would open the door to resolving the tension

292 between overall accuracy, limited by of the number of g-points, and computational cost.

293 Similarly, using spectrally sparse, temporally dense calculations provides a richer set

294 choices in how radiation may be coupled to dynamical models to minimize biases. Both

295 classes of approximations examined here appear to slightly modify the distribution of

296 temperatures, relative to a reference calculation, at quasi-climatological time scales, in

297 contrast to the introduction of sampling noise in cloud properties [*Pincus et al.*, 2006;

298 *Barker et al.*, 2008]. The modification is small but undesirable, and we are seeking ways

299 to reduce the impact of noise. One promising approach is to samples teams in Eq. 4 with

300 replacement, rather than entirely independently at each location and time.

301 The results of section 4, especially Figure 6, suggest that computational effort is the

302 primary determinant of accuracy in coupling these two radiative approximation to short-

303 term forecasting models. This comparison is limited, however, since it does not account

304 for true model errors or how such errors might depend on sampling strategy. Some

305 real-life forecasts errors, such as the damping of surface heating caused by convective

306 clouds forming in response to initial surface heating, may well be due to under-sampled

307 temporal variability. On the other hand, even modest noise in surface fluxes may lead to

308 forecast errors when the coupling between radiation and atmospheric state is strong, as

309 in nocturnal stable boundary layers. Thus it remains to be seen whether the statistical

310 robustness of Equation 4 will translate into improved weather forecasts.

311 Spectrally-sparse, temporally-dense radiation calculations, at least as implemented here,

312 disturb simulations with ECHAM6 at least as much as infrequent broadband calculations

313 of the same computational cost. Time steps in ECHAM are relatively short so the compar-

314 ison may be even worse in models with longer time steps. Other considerations may make

315 Eq. 4 desirable, however, especially the convergence of spectral teams with increasing

316 resolution, more uniform distribution of computation time, and the book-keeping sim-

317 plifications that arise when some estimate of radiation is computed every time step. In

318 ECHAM, for example, shortwave fluxes are computed for all points (using a very small

319 minimum solar zenith angle for nighttime points) so that temporal interpolation across the

320 sunrise boundary is smooth. Replacing infrequent broadband calculations with frequent

321 spectral samples makes this transition smooth (in aggregate), so that shortwave fluxes

322 are not required at nighttime points, which alone represents a substantial computational

323 savings.

324   Though radiation ultimately determines earth's climate the coupling between radiative

325 fluxes and the rest of the atmosphere is loose, such that radiation strongly influences the

326 flow only where its effects can accumulate over time, as occurs in descending branches of

327 the general circulation or at the tops of stratocumulus clouds. The approach to radiation

328 calculations proposed here exploits this loose coupling, trading instantaneous accuracy

329 at infrequent intervals for statistical accuracy with more complete sampling of time vari-

330 ability. By traditional measures of error [e.g. the comparison of instantaneous fluxes to

331 benchmark calculations, as in *Oreopoulos et al.*, 2012], large instantaneous errors make

332 spectral sampling a poor idea. We argue that a more appropriate benchmark is the accu-

333 mulated affect of approximation errors on the solution as a whole. By this more holistic

334 measure of accuracy, frequent but sparse sampling becomes much more attractive because

335 the loose coupling of radiation to the flow means that unbiased solutions with large local

336 errors are more desirable than solutions with small biases, even if their local errors are

337 also small. The present work demonstrates a new path towards accuracy that, in some

338 situations, may converge more quickly to the desired solution.

References

348 Barker, H. W., J. N. S. Cole, J.-J. Morcrette, R. Pincus, P. Raeisaenen, K. von Salzen,

349 and P. A. Vaillancourt, The Monte Carlo Independent Column Approximation: An

350 assessment using several global atmospheric models, *Quart. J. Royal Met. Soc.*, *134* (635,

351 Part B), 1463–1478, doi:10.1002/qj.303, 2008.

352 Buizza, R., M. Miller, and T. N. Palmer, Stochastic representation of model uncertainties

353 in the ECMWF ensemble prediction system, *Quart. J. Royal Met. Soc.*, *125* (560), 2887–

354 2908, doi:10.1002/qj.49712556006, 1999.

355 Chevallier, F., F. Chéruy, N. A. Scott, and A. Chédin, A Neural Network Approach for

356 a Fast and Accurate Computation of a Longwave Radiative Budget, *J. Appl. Meteor.*,

357 *37* (11), 1385–1397, doi:10.1175/1520-0450(1998)037<1385:ANNAFA>2.0.CO;2, 1998.

358 Eckermann, S. D., Explicitly Stochastic Parameterization of Nonorographic Gravity Wave

359 Drag, *J. Atmos. Sci.*, *68* (8), 1749–1765, doi:10.1175/2011JAS3684.1, 2012.

360 Evans, K. F., and A. Marshak, Numerical Methods, in *3D Radiative Transfer in Cloudy*

361 *Atmospheres*, edited by A. Marshak and A. B. Davis, pp. 243–282, Springer, Berlin

362 Heidelbeg New York, 2005.

363 Fu, Q., and K. N. Liou, On the correlated k-distribution method for radiative trans-

364 fer in nonhomogeneous atmospheres, *J. Atmos. Sci.*, *49*, 2139–2156, doi:10.1175/1520-

365 0469(1992)049<2139:OTCDMF>2.0.CO;2, 1992.

366 Hill, P. G., J. Manners, and J. C. Petch, Reducing noise associated with the Monte Carlo

367 Independent Column Approximation for weather forecasting models, *Quart. J. Royal*

368 *Met. Soc.*, *147* (654), 219–228, doi:10.1002/qj.732, 2011.

369   Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and

370   W. D. Collins, Radiative forcing by long-lived greenhouse gases: Calculations with

371   the AER radiative transfer models, *J. Geophys. Res.*, *113*(D13), D13,103, doi:

372   10.1029/2008JD009944, 2008.

373   Kinne, S., D. ODonnel, P. Stier, S. Kloster, K. Zhang, H. Schmidt, S. Rast, M. Giorgetta,

374   T. Eck, and B. Stevens, A new global aerosol climatology for climate studies, *J. Adv.*

375   *Model. Earth Syst.*, 2012.

376   Krasnopolsky, V. M., M. S. Fox-Rabinovitz, and A. A. Belochitski, Decadal Cli-

377   mate Simulations Using Accurate and Fast Neural Network Emulation of Full,

378   Longwave and Shortwave, Radiation, *Mon. Wea. Rev.*, *136*(10), 3683–3695, doi:

379   10.1175/2008MWR2385.1, 2008.

380   Lacis, A. A., and V. Oinas, A description of the correlated k-distribution method for

381   modeling non-grey gaseous absorption, thermal emission, and multiple scattering in

382   vertically inhomogeneous atmospheres, *J. Geophys. Res.*, *96*, 9027–9063, 1991.

383   Lott, F., L. Guez, and P. Maury, A stochastic parameterization of non-orographic gravity

384   waves: Formalism and impact on the equatorial stratosphere, *Geophys. Res. Lett.*, *39*(6),

385   L06,807, doi:10.1029/2012GL051001, 2012.

386   Manners, J., J.-C. Thelen, J. Petch, P. Hill, and J. M. Edwards, Two fast radiative

387   transfer methods to improve the temporal sampling of clouds in numerical weather

388   prediction and climate models, *Quart. J. Royal Met. Soc.*, *135*(639, Part B), 457–468,

389   doi:10.1002/qj.385, 2009.

390   Mauritsen, T., B. Stevens, E. Roeckner, T. Crueger, M. Esch, M. Giorgetta, H. Haak,

391   J. Jungclaus, D. Klocke, D. Matei, U. Mikolajewicz, D. Notz, R. Pincus, H. Schmidt,

392   and L. Tomassini, Tuning the climate of a global model, *J. Adv. Model. Earth Syst.*, *4*,

393   M00A01, doi:10.1029/2012MS000154, 2012.

394   Meador, W. E., and W. R. Weaver, Two-Stream Approximations to Radiative

395   Transfer in Planetary Atmospheres: A Unified Description of Existing Meth-

396   ods and a New Improvement, *J. Atmos. Sci.*, *37*(3), 630–643, doi:10.1175/1520-

397   0469(1980)037<0630:TSATRT>2.0.CO;2, 1980.

398   Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, RRTM, a

399   validated correlated-k model for the longwave, *J. Geophys. Res.*, *103*, 16,663–16,682,

400   doi:10.1029/97JD00237, 1997.

401   Mlawer, E. J., P. D. Brown, S. A. Clough, L. C. Harrison, J. J. Michalsky, P. W. Kiedron,

402   and T. Shippert, Comparison of spectral direct and diffuse solar irradiance measure-

403   ments and calculations for cloud-free conditions, *Geophys. Res. Lett.*, *27*(17), 2653–

404   2656, doi:10.1029/2000GL011498, 2000.

405   Morcrette, J.-J., On the Effects of the Temporal and Spatial Sampling of Radiation

406   Fields on the ECMWF Forecasts and Analyses, *Mon. Wea. Rev.*, *128*(3), 876–887,

407   doi:10.1175/1520-0493(2000)128<0876:OTEOTT>2.0.CO;2, 2000.

408   Morcrette, J.-J., H. W. Barker, J. N. S. Cole, M. J. Iacono, and R. Pincus, Impact of a

409   New Radiation Package, McRad, in the ECMWF Integrated Forecasting System, *Mon.*

410   *Wea. Rev.*, *136*(12), 4773–4798, doi:10.1175/2008MWR2363.1, 2008.

411   Oreopoulos, L., and H. W. Barker, Accounting for subgrid-scale cloud variability in a

412   multi-layer 1D solar radiative transfer algorithm, *Quart. J. Royal Met. Soc.*, *125*(553,

413   Part A), 301–330, doi:10.1002/qj.49712555316, 1999.

414 Oreopoulos, L., E. Mlawer, J. Delamere, T. Shippert, J. N. S. Cole, B. Fomin, M. J.

415   Iacono, Z. Jin, J. Li, J. C. Manners, P. Räisänen, F. Rose, Y. Zhang, M. J. Wilson,

416   and W. B. Rossow, The Continual Intercomparison of Radiation Codes: Results from

417   Phase I, *J. Geophys. Res.*, *117*(D6), D06,118, doi:10.1029/2011JD016821, 2012.

418 Pauluis, O., and K. Emanuel, Numerical instability resulting from infrequent cal-

419   culation of radiative heating, *Mon. Wea. Rev.*, *132*, 673–686, doi:10.1175/1520-

420   0493(2004)132<0673:NIRFIC>2.0.CO;2, 2004.

421 Pincus, R., and B. Stevens, Monte Carlo Spectral Integration: a consistent approximation

422   for radiative transfer in large eddy simulations, *J. Adv. Model. Earth Syst.*, *1*, doi:

423   10.3894/JAMES.2009.1.1, 2009.

424 Pincus, R., H. W. Barker, and J.-J. Morcrette, A fast, flexible, approximate technique

425   for computing radiative transfer in inhomogeneous cloud fields, *J. Geophys. Res.*, *108*,

426   4376, doi:10.1029/2002JD003322, 2003.

427 Pincus, R., R. S. Hemler, and S. A. Klein, Using Stochastically Generated Subcolumns

428   to Represent Cloud Structure in a Large-Scale Model, *Mon. Wea. Rev.*, *134*(12), 3644–

429   3656, doi:10.1175/MWR3257.1, 2006.

430 Räisänen, P., and H. W. Barker, Evaluation and optimization of sampling errors for the

431   Monte Carlo Independent Column Approximation, *Quart. J. Royal Met. Soc.*, *130*(601),

432   2069–2085, doi:10.1256/qj.03.215, 2004.

433 Räisänen, P., H. W. Barker, M. F. Khairoutdinov, J. Li, and D. A. Randall, Stochastic

434   generation of subgrid-scale cloudy columns for large-scale models, *Quart. J. Royal Met.*

435   *Soc.*, *130*(601), 2047–2067, doi:10.1256/qj.03.99, 2004.

436   Stevens, B., M. Giorgetta, E. Monika, T. Mauritsen, T. Crueger, S. Rast, M. Salz-

437   mann, H. Schmidt, J. Bader, K. Block, R. Brokopf, I. Fast, S. Kinne, L. Kornblueh,

438   U. Lohmann, R. Pincus, T. Reichler, and E. Roeckner, The Atmospheric Component

439   of the MPI-M Earth System Model: ECHAM6, *J. Adv. Model. Earth Syst.*, 2012.

440   Strimmer, K., A unified approach to false discovery rate estimation, *BMC Bioinformatics*,

441   *9*(1), 303, doi:10.1186/1471-2105-9-303, 2008a.

442   Strimmer, K., fdrtool: a versatile R package for estimating local and tail area-based false

443   discovery rates, *Bioinformatics*, *24*(12), 1461–1462, 2008b.

444   Taylor, K. E., R. J. Stouffer, and G. A. Meehl, An Overview of CMIP5 and the Experiment

445   Design, *Bull. Amer. Meteor. Soc.*, *93*(4), 485–498, doi:10.1175/BAMS-D-11-00094.1,

446   2012.

447   Tompkins, A. M., A prognostic parameterization for the subgrid-scale variability of water

448   vapor and clouds in large-scale models and its use to diagnose cloud cover, *J. Atmos.*

449   *Sci.*, *59*(12), 1917–1942, doi:10.1175/1520-0469(2002)059<1917:APPFTS>2.0.CO;2,

450   2002.

451   Turner, D. D., D. C. Tobin, S. A. Clough, P. D. Brown, R. G. Ellingson, E. J. Mlawer, R. O.

452   Knuteson, H. E. Revercomb, T. R. Shippert, W. L. Smith, and M. W. Shephard, The

453   QME AERI LBLRTM: A Closure Experiment for Downwelling High Spectral Resolution

454   Infrared Radiance, *J. Atmos. Sci.*, *61*(22), 2657–2675, doi:10.1175/JAS3300.1, 2004.

455   Venema, V., A. Schomburg, F. Ament, and C. Simmer, Two adaptive radiative transfer

456   schemes for numerical weather prediction models, *Atmos. Chem. Phys.*, *7*(21), 5659–

457   5674, doi:10.5194/acp-7-5659-2007, 2007.

Wilks, D. S., On "Field Significance" and the False Discovery Rate, *J. Appl. Meteor. Climatol.*, *45*(9), 1181–1189, doi:10.1175/JAM2404.1, 2006.

Xu, K. M., and D. A. Randall, Impact of Interactive Radiative Transfer on the Macroscopic Behavior of Cumulus Ensembles. Part I: Radiation Parameterization and Sensitivity Tests, *J. Atmos. Sci.*, *52*, 785–799, doi:10.1175/1520-0469(1995)052<0785:IOIRTO>2.0.CO;2, 1995.
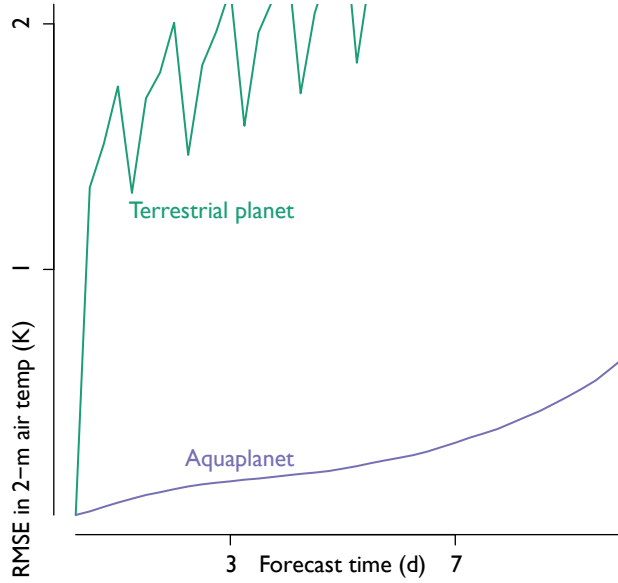
Figure 1.   The impact of randomly sampling a single spectral interval at every time step (Monte Carlo Spectral Integration, Eq. 3). Root-mean-square differences are computed with respect to reference forecasts for an Earth-like planet (green), in which the land temperature responds to surface fluxes, and an aquaplanet with fixed sea surface temperatures (purple). The diurnal variation in RMSE for the earth-like planet tracks the diurnal variation in global 2-m air temperature caused by the asymmetric distribution of land. Sampling noise in radiative heating rates within the atmosphere drives modest differences, while noise can persist and modify the simulation when the surface temperature can change in response to surface radiation fluxes, both because heating rates are substantially larger at the surface than in the atmosphere and because sampling errors at the surface are not mixed by the flow as they are in the atmosphere. This implies that accuracy in radiation parameterizations depends in part on limiting perturbations to downwelling fluxes at the surface.
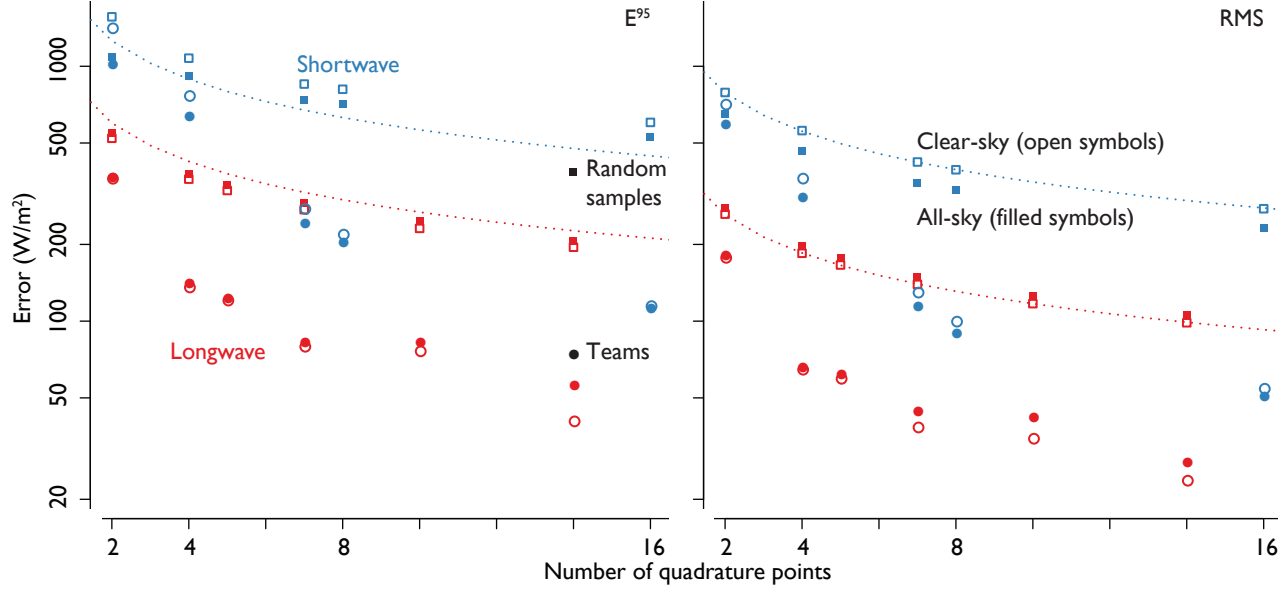
Figure 2.   Errors in shortwave (blue) and longwave (red) surface fluxes for various spectral

sampling strategies as a function of the number of quadrature points used. Errors are accumulated

over roughly 73000 sample columns representing four snapshots from a single day of ECHAM.

Root-mean square error (right panel) decreases as $1/\sqrt{(n)}$ when samples are chosen randomly

(Eq. 3, squares) though E^{95} (left) does not, but both decrease much more quickly using "teams"

of g-points chosen to minimize E^{95} (Eq. 4, circles) than as compared to calculations using the

same number of quadrature points, but chosen at random. Errors for all-sky fluxes (closed circles)

are marginally smaller in the shortwave and larger in the longwave than for the clear-skies used

to optimize the team members primarily because clouds decrease SW surface and increase LW

fluxes.
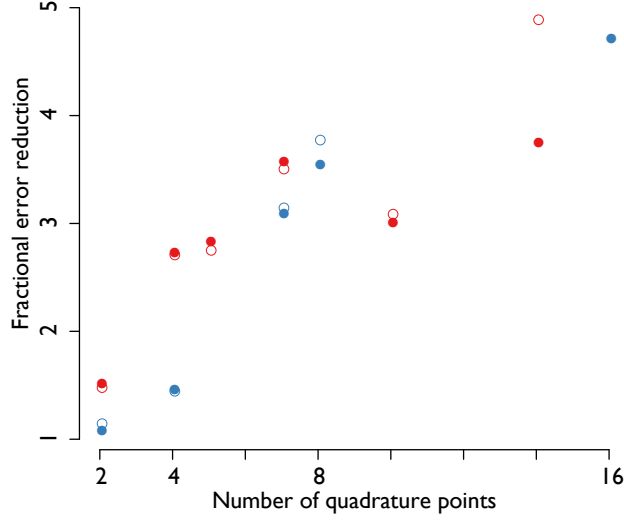
Figure 3.   Ratio of 95-th percentile error E^{95} in downwelling surface flux using teams of spectral intervals, compared to Monte Carlo sampling of spectral intervals, as a function of the number of samples. Improvements for clear-sky fluxes (open circles) and all-sky fluxes (closed circles) are commensurate. At small sample sizes error in the longwave fluxes (red) is reduced more dramatically by using teams than are shortwave fluxes (blue). The error reduction provided by teams increases with the number of samples but even teams of modest sizes (~ 8 members) provide accuracy comparable to ~ 10 times as many random samples.
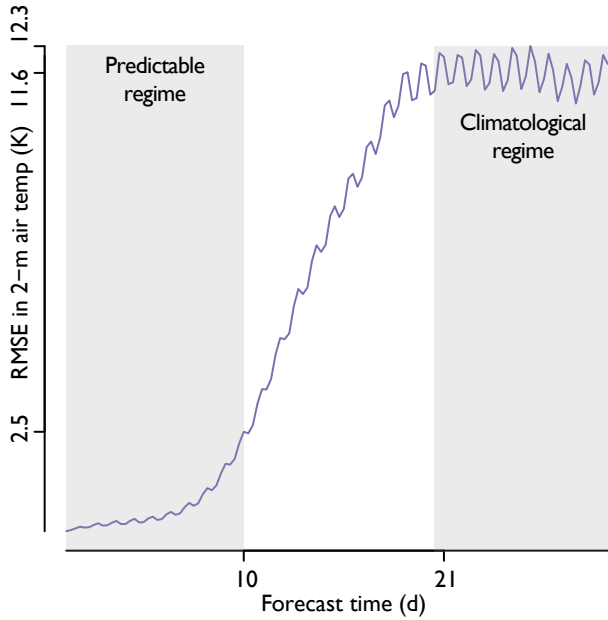
Figure 4.   Root-mean-square difference in 2 m air temperature, relative to a reference forecast, as a function of forecast lead time, for a simulation making broadband radiation calculations every 2 h while all other physical processes use a 7.5 min time step. The forecasts diverge slowly during the first ten days (the weather forecasting regime, roughly), so that approximations can be judged by computing point-by-point root-mean-square differences with the reference. After roughly three weeks the forecasts are not predictable in a deterministic sense, so parameterization accuracy is judged by comparing distributions.
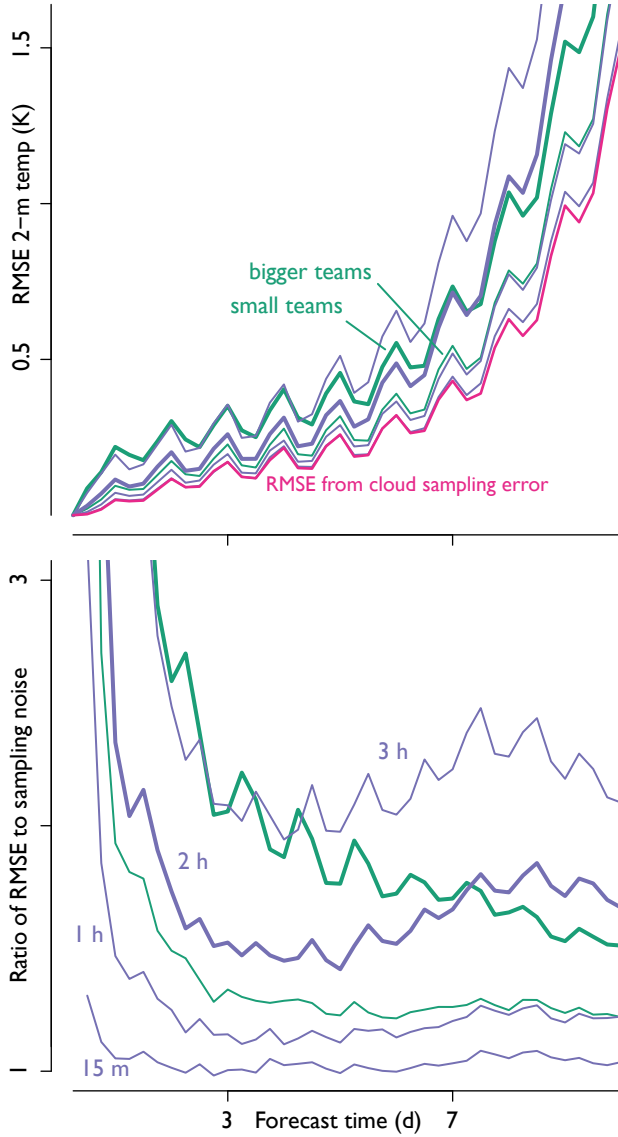
Figure 5.   Top: Root-mean-square difference in 2 m air temperature, relative to a reference forecast, as a function of forecast lead time for two methods of coupling radiation and dynamics in a global model. Broadband computations applied sparsely in time at intervals ranging from 3 h to 15 min are shown in purple and two examples of "teams" of spectral intervals (as described in the text, using (7, 16) points in the shortwave and (4, 10) in the longwave) are in green. Differences between independent realizations of the reference forecast (in pink) show the differences to be expected from sampling errors in cloud state introduced by the Monte Carlo Independent Pixel Approximation (McICA). Bottom: Ratio of RMS difference of different coupling approximations to the RMS difference due to finite sampling of cloud states. All approximations change model evolution; larger approximations (fewer team members or less frequent sampling in time) degrade the sampling of the diurnal cycle.
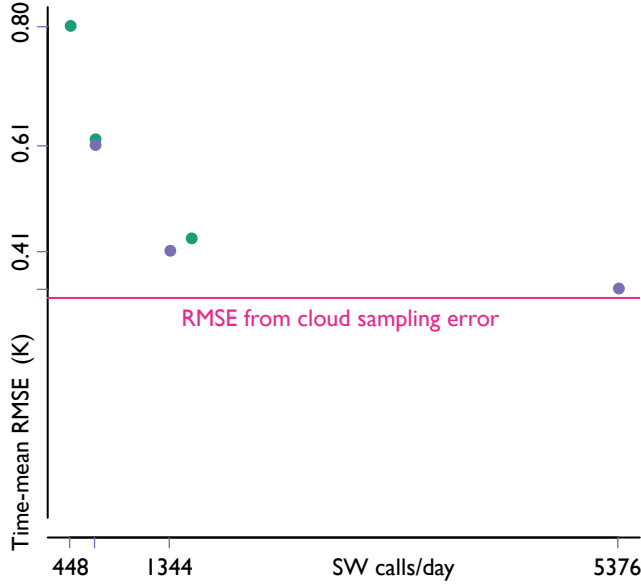
Figure 6. Time-mean error versus a rough measure of computational cost for two methods of coupling radiation and dynamics in a global model. Broadband computations applied sparsely in time are shown in purple and the "teams" of spectral intervals described in the text are in green. Error is measured as the mean over ten days of the global root-mean-square difference in 2 m air temperature relative to a reference forecast, and computational cost as the daily number of calls to the shortwave solver (which dominates the overall cost). The pink line shows the minimum achievable error (i.e. the error introduced by another realization of cloud states sampled in the Monte Carlo Independent Pixel Approximation). Approximation errors for the two sampling strategies are commensurate for a given computational cost.

Table 1.   Time mean fraction of of statistically insignificant η_0 differences in 2-m air temperature between a reference calculation and various approximations for coupling radiation to a global model. The first three approximations are temporally-dense, spectrally-sparse calculations using Equation 4 (the first is the special case using Equation 3). The latter four make spectrally-dense broadband calculations at specified time intervals. All approximations change the simulation of 2-m air temperature by detectable amounts; for a given computational effort, frequent use of spectral teams introduces slightly larger changes than less frequent broadband calculations.

Approx.	SW calls/day	$\overline{\eta_0}$
$m_{\mathrm{SW}} = 1, m_{\mathrm{LW}} = 1$	192	0.804
$m_{\mathrm{SW}} = 7, m_{\mathrm{LW}} = 4$	1344	0.943
$m_{\mathrm{SW}} = 16, m_{\mathrm{LW}} = 10$	3072	0.977
$T = 15$ m	10752	0.971
$T = 1$ h	2688	0.985
$T = 2$ h	1344	0.955
$T = 3$ h	896	0.963